PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12P 21/02	A1	(11) International Publication Number: WO 92/16642
		(43) International Publication Date: 1 October 1992 (01.10.92)

(21) International Application Number: PCT/US92/01598

(22) International Filing Date: 26 February 1992 (26.02.92)

(30) Priority data:
671,376 19 March 1991 (19.03.91) US

(71) Applicant: OMNIGENE, INC. [US/US]; 85 Bolton Street, Cambridge, MA 02140 (US).

(72) Inventors: SLOMA, Alan ; 20 James Street, Watertown, MA 02172 (US). RUFO, Gerald, A., Jr. ; 6 Baron Park Lane, Apt. 30, Burlington, MA 01803 (US). PERO, Janice ; 20 Solomon Pierce Road, Lexington, MA 02173 (US).

(74) Agent: CLARK, Paul, T.; Fish & Richardson, 225 Franklin Street, Boston, MA 02110-2804 (US).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), MC (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: RESIDUAL PROTEASE-III

(57) Abstract

A *Bacillus* cell containing a mutation in the residual protease III (rp-III) gene resulting in the inhibition of the production by the cell of proteolytically active RP-III.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FI**	Finland	**ML**	Mali
AU	Australia	**FR**	France	**MN**	Mongolia
BB	Barbados	**GA**	Gabon	**MR**	Mauritania
BE	Belgium	**GB**	United Kingdom	**MW**	Malawi
BF	Burkina Faso	**GN**	Guinea	**NL**	Netherlands
BG	Bulgaria	**GR**	Greece	**NO**	Norway
BJ	Benin	**HU**	Hungary	**PL**	Poland
BR	Brazil	**IE**	Ireland	**RO**	Romania
CA	Canada	**IT**	Italy	**RU**	Russian Federation
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		
ES	Spain	**MG**	Madagascar		

Residual Protease-III

Background of the Invention

This invention relates to Bacillus strains useful for the expression and secretion of desired polypeptides (as used herein, "polypeptide" means any useful chain of amino acids, including proteins).

Bacillus strains have been used as hosts to express heterologous polypeptides from genetically engineered vectors. The use of a Gram positive host such as Bacillus avoids some of the problems associated with expressing heterologous genes in Gram negative organisms such as E. coli. For example, Gram negative organisms produce endotoxins which may be difficult to separate from a desired product. Furthermore, Gram negative organisms such as E. coli are not easily adapted for the secretion of foreign products, and the recovery of products sequestered within the cells is time consuming, tedious, and potentially problematic. In addition, Bacillus strains are non-pathogenic and are capable of secreting proteins by well-characterized mechanisms.

A general problem in using Bacillus host strains in expression systems is that they produce large amounts of proteases which can degrade heterologous polypeptides before they can be recovered from the culture media. The production of the majority of these proteases occurs at the end of the exponential growth phase. At this time, conditions of nutrient deprivation exist and the cells are preparing for sporulation. The two major extracellular proteases are an alkaline serine protease (subtilisin), the product of the apr gene, and a neutral metalloprotease, the product of the npr gene. Secretion of these proteases occurs into the medium, whereas the major intracellular

serine protease, Isp-I, is produced within the cells. Other investigators have created genetically altered Bacillus strains that produce below normal levels of one or more of these three proteases. These strains still produce high enough levels of protease to cause the degradation of heterologous gene products prior to purification.

Stahl et al. (J. Bact., 1984, 158:411) disclose a Bacillus protease mutant in which the chromosomal subtilisin structural gene was replaced with an in vitro derived deletion mutation. Strains carrying this mutation had only 10% of the wild-type extracellular production of serine protease activity. Yang et al. (J. Bact., 1984, 160:15) disclose a Bacillus protease mutant in which the chromosomal neutral protease gene was replaced with a gene having an in vitro derived deletion mutation. Fahnestock et al. (WO 86/01825) describe the construction of Bacillus strains lacking subtilisin activity by replacing the native chromosomal gene sequence with a partially homologous DNA sequence containing an inserted inactivating segment. Kawamura et al. (J. Bact., 1984, 160:442) disclose Bacillus strains carrying lesions in the npr and apr genes. These strains express less than 4% of the extracellular protease activity levels observed in wild-type strains. Koide et al. (J. Bact., 1986, 167:110) disclose the cloning and sequencing of the isp-1 gene and the construction of an Isp-1 negative mutant by chromosomal integration of an artificially deleted gene.

Sloma et al., 1990 J. Bact. 172:1024-1029, employed B. subtilis deleted for the three major proteases (apr, npr, isp) in order to identify three additional residual proteases (epr, bpr, mpr). Blackburn et al., WO 89/10976 also used sporulation competent apr-, npr- strains to isolate what they alledge to be a residual serine protease

(rsp) which lacks amino terminal homology to known bacillus proteases.

Genetically altered strains which are deleted for both the major extracellular protease genes (apr and npr) and three residual protease genes (epr, bpr, mpr) produce virtually undetectable levels of protease activity in standard protease assays. However, a resorufin-labeled casein substrate, can be used to detect minor protease activities which are responsible for degradation of some heterologous polypeptides and proteins.

Summary of the Invention

The invention provides a novel protease, RP-III, and improved Bacillus cells containing mutations in the previously uncharacterized RP-III encoding gene (vpr); the cells also preferably contain mutations in the one or more or any combination of extracellular protease encoding apr, npr, epr, bpr, and mpr genes, resulting in the inhibition by the cells of production of these proteases. The bpr and mpr genes are also known as rp-I and rp-II, respectively.

Preferably, the mutation of the invention involves a mutation in the rp-III gene (recently named vpr) which inhibits the production by the cell of the proteolytically active RP-III. (As used herein, mutation can refer to a deletion within or of the coding region of a gene, a substitution of one or more base pairs for one or more naturally occurring base pairs, or an insertion of one or more base pairs within the coding region of a gene.) Most preferably, the mutation of the invention is a deletion within the coding region of the gene, including deletion of the entire coding region; alternatively, the mutation can consist of a substitution of one or more base pairs for naturally ocurring base pairs, or an insertion within the protease coding region.

The Bacillus cells of the invention may also contain a mutation in the isp-1 gene encoding intracellular serine protease I and may, in addition, contain a mutation which blocks sporulation and thus reduces the cell's capacity to produce sporulation dependent proteases; preferably, this mutation blocks sporulation at an early stage, most preferably, this mutation is the spoOA mutation (described below). The invention further provides a method for producing stable heterologous polypeptides in a Bacillus host cell by modifying the host to contain mutations in the apr, npr, and rp-III genes and in one or more of the genes including the epr gene, the bpr gene, and the mpr (rp-II) gene. The method may include introducing into the Bacillus host cell a gene encoding a heterologous polypeptide that is modified so as to be expressed in the Bacillus host; such gene modifications may include but are not limited to a compatible promoter sequence, enhancer sequence, and/or ribosome binding site.

The invention also features purified DNA, expression vectors containing DNA, and host Bacillus cells transformed with DNA encoding RP-III; preferably, such DNA is derived from Bacillus subtilis.

The invention also features the isolation of a substantially pure previously uncharacterized residual protease (RP-III); as used herein, "substantially pure" means greater than 90% pure by weight.

The term "rp-III gene" herein means the respective gene corresponding to this designation in Bacillus subtilis, and the evolutionary homologues of this gene in other Bacillus species, which homologues, as is the case for other Bacillus proteins, can be expected to vary in minor respects from species to species. In many cases, sequence homology between evolutionary homologues is great enough so that a

gene derived from one species can be used as a hybridization probe to obtain the evolutionary homologue from another species, using standard techniques. In addition, of course, those terms also include genes in which base changes have been made which, because of the redundancy of the genetic code, do not change the encoded amino acid residue or which produce conservative changes (to an amino acid of similar hydrophobicity or charge distribution) to a few amino acids.

Using the procedures described herein, we have produced Bacillus strains which are significantly reduced in their ability to produce proteases, and are therefore useful as hosts for the expression, without significant degradation, of heterologous polypeptides capable of being secreted into the culture medium. We have found that the Bacillus cells of the invention, even though containing several mutations in genes encoding related activities, are not only viable but healthy.

Any desired polypeptide can be expressed according to the invention, e.g., medically useful proteins such as hormones, vaccines, antiviral proteins, antitumor proteins, antibodies or clotting proteins; and agriculturally and industrially useful proteins such as enzymes or pesticides, and any other polypeptide that is normally degraded by RP-III.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Description of Preferred Embodiments

The drawings will first be briefly described.

Drawings

Fig. 1 is a comparison of N-terminal sequence of RP-III to a composite N-terminal sequence deduced from known

B. subtilis serine protease sequences encoded by apr, epr, bpr and isp-1.

Fig. 2 is the N-terminal sequence of RP-III and corresponding sequence of the "guess-mer" oligonucleotide probe used to identify the rp-III gene.

Fig. 3 is a restriction map of a DNA fragment containing the rp-III coding region and shows approximate locations of rp-III subclones.

Fig. 4 is the DNA sequence of DNA encoding the rp-III gene.

General Strategy for Creating Protease Deficient Bacillus Strains

General Methods

In order to detect residual protease activity remaining in B. subtilis after removal of other known proteases, a strain must be constructed which lacks the known proteases. A Bacillus strain which is substantially devoid of extracellular proteolytic activity is described in EPA 0 369 817 A2, by Sloma et al., hereby incorporated by reference. A similar strain which contains multiple mutations which inactivate apr, npr, isp-1, epr, bpr, and mpr was prepared and assayed for residual serine protease activity using resorufin-labeled casein (Boehringer-Mannheim) as a substrate. Residual serine protease RP-III was detected in the multiply mutated strain; its activity was monitored throughout purification using the same substrate. The purification and characterization of RP-III and isolation of the gene encoding RP-III are described below, along with a procedure for generating a Bacillus strain containing a mutation which inactivates the RP-III-encoding gene.

General Methods

Construction of a multiply-mutated Bacillus strain is described by Sloma et al EPA 0 369 817 A2. Isolation of B. subtilis chromosomal DNA was as described by Dubnau et al., (1971, J. Mol. Biol., 56: 209). B. subtilis strains were grown on tryptose blood agar base (TBAB) (Difco Laboratories) or minimal glucose medium and were made competent by the procedure of Anagnostopoulos et al., (J. Bact., 1961, 81: 741). E. coli JM107 was grown and made competent by the procedure of Hanahan (J. Mol). Biol., 1983, 166: 587). Plasmid DNA from B. subtilis and E. coli were prepared by the lysis method of Birnboim et al. (Nucl. Acid. Res., 1979, 7: 1513). Plasmid DNA transformation in B. subtilis was performed as described by Gryczan et al., (J. Bact., 1978, 134: 138).

Protease assays

Resorufin-labelled casein or ^{14}C-casein was used for RP-III assays. Culture supernatant samples were assayed either 2 or 20 hours into stationary phase. Inhibitors were pre-incubated with the supernatant for 30 minutes at room temperature. Where a very small amount of residual protease activity was to be measured, ^{14}C-casein or resorufin-labelled casein was used as the substrate.

In the ^{14}C-casein test, culture supernatant (100 μl) was added to 100 μl of 50mM Tris, 5mM $CaCl_2$, pH 8, containing 1 X 10^5 cpm of ^{14}C casein (New England Nuclear). The solutions were incubated at 37° C for 30 minutes. The reactions were then placed on ice and 20 μg of BSA were added as carrier protein. Cold 10% TCA (600 μl) was added and the mix was kept on ice for 10 minutes. The solutions were centrifuged to spin out the precipitated protein and the supernatants counted in a scintillation counter.

The resorufin-labeled casein assay involved incubation of culture supernatant with an equal volume of resorufin-labelled casein in 50 mM Tris, 5mM $CaCl_2$, pH 8.0, at 45° C for 1 hour. Following incubation, unhydrolyzed substrate was precipitated with TCA and centrifuged. The supernatant (400ml) was made alkaline with 500mM Tris (pH 8.8) and the resulting chromogenic supernatant was quantitated spectrophotometrically at 574 nm.

Parental Strains

A number of Bacillus strains were used as sources for strains of the current invention.

Strain GP216, containing deletions within the four protease genes (apr, npr, isp-1, and epr), and strain GP240, containing deletions with the five protease genes (apr, npr, isp-1, epr, and bpr (rp-I)), were prepared as described by Sloma et al., EPA 0 369 817 A2. Strain GP241, isogenic to GP240 except for the hpr gene, was constructed from strain GP216 by transformation of GP216 with a plasmid (pUC derivative called pJMhpr2, Perego and Hoch, J. Bacteriology 170:2560, 1988) containing a mutated hpr gene and a *cat* gene. hpr encodes a repressor of protease production in *Bacillus*. GP216 was transformed with pJMhpr2 and transformants were selected on chloramphenicol. Chromosomal DNA was extracted from chloramphenicol resistant colonies and analyzed by Southern hybridization. One clone was recovered which contained two copies of the hpr-2 gene resulting from a double crossover between homologous sequences on the vector and in the chromosome. The clone was grown in the absence of drug selection, and one chloramphenicol sensitive colony was designated BI114. Strain GP241 was constructed by introducing the deleted bpr (rp-I) gene into BI114 using the plasmid pKT3 in the same

manner as described in Sloma et al. (EPA 0 369 817 A2) for the introduction of the deleted bpr (rp-I) gene into GP216 generating GP240.

Strain GP263, carrying a mutation in mpr was prepared from GP241 as follows. Plasmid pCR125, carrying the phleomycin resistance gene inserted in a deleted mpr gene (Sloma et al., EPA 0 369 817 A2), was digested with EcoR1 and the linear plasmid DNA was used to transform GP241 to phleomycin resistance. Resistant transformants were selected by plating the transformed cells onto TBAB plates containing a gradient of 0-5 μg/ml phleomycin across the plate. Transformants that were resistant to approximately 2.5 ug/ml phleomycin on the plates were single colony purified on TBAB phleomycin plates and thereafter grown on TBAB without selective antibiotic. One transformant isolated following this treatment was designated GP263.

GP263 was used to generate two additional strains, GP264 and GP275. GP264 has the sacQ* regulatory element chromosomally integrated via transformation with the plasmid pDP104, as described by Sloma et al., EPA 86308356.4. GP275 was produced by fully deleting the already-inactivated mpr (rp-II) gene from GP263. Since inactivation of mpr was due to an insertion of the phleomycin resistance gene into mpr, the deletion of mpr was accomplished by transformation of GP263 with a plasmid containing a deleted mpr and chloramphenicol resistance genes in contiguous array. Transformants were selected on chloramphenicol. Isolated colonies were then grown in the absence of selection and replica plated. GP275 was isolated as both choloramphenicol and phleomycin sensitive.

Identification of A Novel Proteolytic Activity

Extracellular protease levels are reduced in culture supernatants of Bacillus strains that do not express the proteases encoded by the six non-essential protease genes, apr, npr, isp-1 epr, bpr and mpr. When these deletions are present in a Spo+ host, there is an approximate 99% reduction in extracellular protease levels compared to the wild-type strain. In order to efficiently produce protease labile products in Bacillus, it is desirable to decrease or eliminate the remaining 1% residual protease activity.

Using the resorufin-labeled casein assay, a novel protease has been identified which is a major component of the residual activity in GP264. This protease may be classified as a serine protease by virtue of its quantitative inhibition by phenylmethylsulfonyl fluoride.

Isolation and Characterization of RP-III

A simple and efficient purification scheme was developed for the isolation of the RP-III protease from spent culture fluids. Cultures were grown in modified MRS lactobacillus media (Difco, with maltose substituted for glucose) and concentrated approximately 20-fold using an Amicon CH2PR system equipped with a S1Y10 spiral cartridge and dialyzed in place against 50mM MES pH 5.5, and allowed to incubate overnight at 0-4°C. The concentrated, crude supernatant containing precipitated protein was centrifuged (Sorvall GSA rotor, 9000 rpm, 30 minutes) and the resulting pellet containing 80-100% of the RP-III protease activity was resuspended in 100 mM Tris, pH 8. The reconstituted pellet was then applied to a 500 ml Superflo (Sepragen) column packed with Q-Sepharose (Pharmacia) equilibrated with 100mM Tris, pH 8. Bound protein containing the RP-III protease was recovered from the column with a 50mM MES, 2.5 M KCl, pH 5.5, step elution.

The high-salt fractions containing protease activity were pooled, concentrated and dialyzed against 50mM MOPS, pH 7, then applied to a 250 ml Superflo column of benzamidine Sepharose (Pharmacia) affinity resin equilibrated with the same buffer. Bound RP-III protease was eluted from the resin with a step of 50mM MOPS, 1 M KCl, pH 7. Proteolytically active high-salt fractions containing RP-III protease were pooled, concentrated and subjected to HPLC size-exclusion chromatography over a semi-preparative SW3000 column equilibrated with 50mM MES, 200mM KCl, pH 6.8. Protease activity was found exclusively in the void volume indicating the RP-III protease exists as part of a large aggregate. Finally, the size-excluded RP-III pool was concentrated, dialyzed against 20mM sodium phosphate, 1M NaCl, 1mM imidazole, pH 7.5, and fractionated over a Progel-TSK chelate-5PW HPLC column charged with Cu++. Activity was eluted with a linear gradient of imidazole to 20mM.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) revealed that the final pool of RP-III protease contained three major Coomassie-staining bands: one at 38.4 kDa and a doublet at 28.5 and 27.1 kDa. Each of these bands were electrophoretically transferred to and cut out of a sheet of PVDF membrane and subjected to amino-terminal sequence analysis. The sequence of the 28.5 kDa protein bore remarkable homology (81%) to a composite sequence of four other B. subtilis serine proteases (apr, subtilisin; epr, extracellular protease; bpr, Bacillopeptidase F, and isp-1, intracellular protease 1) as well as to Bacillopeptidase F itself (65% homology). The proteolytic activity in this band is referred to herein as RP-III. Figure 1 illustrates the amino-terminal sequence of RP-III and its comparison to a composite sequence deduced

from the amino acid sequences of the aformentioned B. subtilis serine proteases.

All five proteases contain six identical residues spaced exactly the same within the N-termini, including the putative active center aspartic acid residue. Sequence analysis of the 27.1kDa lower band revealed it is most likely a proteolytic fragment of the 28.4kDa upper band since both proteins have identical amino-terminal sequences from residue 10 to residue 29. The loss of residues 1-9 on the lower 27.1kDa band accounts for its faster mobility on SDS-PAGE compared to the upper 28.4kDa band.

Figure 2 shows the amino-terminal sequence obtained from RP-III and the sequence of the oligomeric probe constructed to identify the gene that codes for RP-III.

Cloning and Sequencing of the rp-III Gene.

Genomic DNA was prepared from Bacillus subtilis GP275, and 10 μg were exhaustively digested with HindIII and probed with the guess-mer shown in Fig 2. The probe hybridized to a 1kb fragment of HindIII-digested genomic DNA; therefore, a 1kb genomic library was prepared from size-selected fragments of 0.8-1.5 kb, using pUC19 as the vector. A clone carrying the rp-III gene was identified in the 1 kb library using standard hybridization techniques (Sambrook et al., 1989, Molecular Cloning, Cold Spring Harbor, NY) and the guess-mer probe shown in Fig. 2. The plasmid isolated from this clone was designated pLLP1.

Southern blot analysis was used to determine the location of useful restriction sites with the rp-III gene (Fig. 3). Southern blots were performed using restriction digests of genomic DNA from GP275 and a probe encompassing the 1kb HindIII fragment from pLLP1. These results led to the preparation of size-selected EcoR1, EcoR1/BglII, EcoR1/HindIII and BglII libraries from GP275 genomic DNA.

Libraries yielding useful clones were prepared in either pIC20H or in pUC19 vectors digested with the apropriate restriction enzymes. pLLP4 and pLLP5 were isolated from 3kb and 0.5-0.8kb EcoR1/BglII pIC20H libraries, respectively, by screening with the 1kb HindIII fragment of pLLP1. pLLP8 was isolated from a 0.5-0.8kb EcoR1/HindIII pUC19 library by screening with the 630 bp BglII fragment of pLLP5.

These clones were used to construct a restricion map of the rp-III gene, after the regions flanking the 1kb HindIII fragment were identified. The DNA sequence was determined between the 5' BglII site of pLLP5 and approximately 1kb beyond the 3' HindIII site of pLLP4 (Figs. 3 and 4).

An open reading frame was found to extend 2457 nucleotides downstream from the 5' Bg1II site. A putative translation initiation codon was identified (Fig 4, underlined nucleotides 40-42), with an accompanying ribosome binding site (Fig. 4, underlined nucleotides 25-32). The amino terminal sequence of the mature protein corresponding to the sequence in Figure 2, was found at nucleotide 520 and is underlined in Figure 4. From the sequence data of Figure 4, the mature protein encoded by the rp-III gene is expected to contain 646 amino acids. Since the isolated protein has an apparent molecular weight of 28,000 d., this would suggest that rp-III undergoes extensive C-terminal processing or proteolysis.

Location of the rp-III Gene on the B. Subtilis Chromosome

Identification of the chromosomal location of the rp-III gene may be accomplished by standard methods, essentially as described by Sloma et al. EPA 0 369 817 A2, for other protease genes. Briefly, the location of the rp-III gene on the B. subtilis chromosome was mapped by

integrating a drug resistance marker into the chromosome at the site of rp-III and using phage PBS1-mediated transduction to determine the location of the drug resistance gene. A fragment containing a neomycin resistance (neo) gene was cloned into the BglII site within the amino terminal coding region of rp-III, as described below to give plasmid pLLP2 which was used to create GP279. Southern blotting techniques and hybridization were used to confirm that the neo gene had integrated into the chromosome, interrupting the rp-III gene. Mapping experiments were then used to indicate that the inserted neo gene and rp-III are linked to the known Bacillus genetic locations, sacA, ctr, and epr, by PBS1 transduction.

Inactivation of the rp-III gene

It is often useful to inactivate the production of functional RP-III protease in microorganisms, particularly when a desired protein is being produced which is succeptible to RP-III proteolysis. The rp-III gene sequence provided herein allows for elimination of RP-III activity by any number of standard methods; including inactivation by insertion of nucleotide sequences into the gene, or by deletion of part or all of the native gene. In general, homologous recombinant techniques may be employed; for example, see Sloma et al. EPA 0 369 817 A2.

The rp-III gene was inactivated by creating an insertion mutation within the native gene. A 2.4kb SmaI to SmaI fragment containing the entire neomycin resistance gene was inserted into the Klenow blunt-ended BglII site of pLLP1, to give the plasmid pLLP2. pLLP2 was then linearized by ScaI digestion and used to transform Bacillus strain GP275. Neomycin resistant strains from this transformation were called GP279 and contained an inactivated rp-III gene. The inactivation of rp-III was confirmed by protease

activity assay, as described above. Strains bearing the insertion mutation were otherwise normal with regard to sporulation and growth.

Heterologous DNA Expression

Cells in which the rp-III gene has been inactivated may be employed to express useful heterologous proteins. Such proteins would typically be of medical, agricultural, or industrial significance. In order to minimize any potential proteolytic damage of the heterologous protein, preferred cells will also be inactivated for apr, npr, epr, bpr, and mpr. Inactivation of additional genes such as isp-1 and spoOA may also be useful.

DNA encoding the desired heterologous proteins must be engineered to contain the proper regulatory sequences including promoter, ribosome binding site, and transcription termination signals. In general, the DNA sequence encoding the protein and its accompanying regulatory sequences must be compatible with expression in the Bacillus host cell of the invention. The introduced DNA containing the expression sequences may reside within the cell in plasmid form or more preferably it may be chromosomally integrated.

The following references are incorporated herein by reference: Guidelines and references for heterologous protein expression and selection of appropriate Bacillus regulatory elements are given in Ganesan et al., 1986 *Bacillus Molecular Genetics and Biotechnology Applications*. Academic press pp. 367-493. Methods useful for the construction of expression vectors are given by Sambrook et al., 1989, *Molecular Cloning a Laboratory Manual* Cold Spring Harbor Laboratory Press.

Other Embodiments

Other embodiments are within the following claims. For example, in some instances it may be desirable to express, rather than mutate or delete, the gene encoding RP-III; for example, to produce the protease for purposes such as improvement of the cleaning activity of laundry detergents or for use in industrial processes. This can be accomplished either by inserting regulatory DNA (any appropriate Bacillus promoter and, if desired, ribosome binding site and/or signal encoding sequence) upstream of the protease-encoding gene or, alternatively, by inserting the protease-encoding gene into a Bacillus expression or secretion vector; the vector can then be transformed into a Bacillus strain for production (or secretion) of the protease, which is then isolated by conventional techniques. Alternatively, the protease can be overproduced by inserting one or more copies of the protease gene on a vector into a host strain containing a regulatory gene such as sacQ*.

Claims

1. A Bacillus cell containing a mutation in the rp-III gene resulting in inhibition of the production by said cell of proteolytically active RP-III.

2. The Bacillus cell of claim 1, further comprising a mutation in each of one or more protease-encoding genes selected from the group: apr, npr, epr, bpr, and mpr, wherein each said mutation results in inhibition of the production by said cell of proteolytically active protease encoded by said gene.

3. The Bacillus cell of claim 2, each said mutation comprising a deletion within the coding region of said gene.

4. The Bacillus cell of claim 3, said cell further containing a mutation in the isp-1 gene encoding an intracellular protease.

5. The Bacillus cell of any of claims 1-4, said cell further containing a mutation which reduces said cell's capacity to produce one or more sporulation-dependent proteases.

6. The Bacillus cell of claim 5 wherein said sporulation-dependent protease mutation blocks sporulation at an early stage.

7. The Bacillus cell of claim 6, said sporulation-blocking mutation being in the spoOA gene.

8. The Bacillus cell of claim 7, said cell being Bacillus subtilis.

9. The Bacillus cell of any one of claims 1-4 and 6-8, further comprising a gene encoding a heterologous polypeptide.

10. The Bacillus cell of claim 5 further comprising a gene encoding a heterologous polypeptide.

12. The cell of claim 9 wherein said heterologous polypeptide is a medically, agriculturally or industrially useful protein.

16. A method for producing a heterologous polypeptide in a Bacillus cell, said method comprising introducing into said cell a gene encoding said heterologous polypeptide, modified to be expressed in said cell, said Bacillus cell containing mutations in the rp-III, apr and npr genes.

17. The method of claim 16 wherein said cell further contains mutations in one or more of the genes, epr, bpr,or mpr.

18. The method of claim 17, said cell further containing a mutation in the isp-1 gene encoding intracellular protease I.

19. The method of claim 16, 17, or 18 wherein said cell further contains a mutation which reduces said cell's capacity to produce one or more sporulation-dependent proteases, said mutation being in the spoOA gene.

20. The method of claim 19 wherein said cell is a Bacillus subtilis cell.

21. The method of claim 20 wherein said heterologous polypeptide is a medically, agriculturally or industrially useful protein.

22. Purified DNA comprising a Bacillus rp-III gene.

23. A vector comprising a Bacillus rp-III gene and regulatory DNA operationally associated with said gene.

24. A Bacillus cell transformed with the vector of claim 23.

25. Substantially pure Bacillus RP-III protease.

26. The DNA of claim 22 wherein said sequence is sequence ID No._____ (Fig. 4).

FIG. 1 - N-TERMINAL AMINO ACID HOMOLOGY BETWEEN RP-III AND OTHER *B. SUBTILIS* SERINE PROTEASES (I.E., BPR, EPR, APR, ISP-I)

RP-III	I	G	A	N	D	A	W	D	L	G	Y	T	G	K	G	I	K	V	A	I	I	D	T	G	V	E
COMPOSITE	I	-	A	-	-	A	W	-	L	G	Y	T	G	K	G	I	K	V	A	-	I	D	T	G	V	E

Δ
Δ
ACTIVE CENTER ASP

COMPOSITE HOMOLOGY - 81%

BPR HOMOLOGY - 65%

FIG. 2 - AMINO-TERMINAL SEQUENCE OF RP-III AND CORRESPONDING "GUESS-MER" PROBE SEQUENCE

	1	2	3	4	5	6	7	8
H_3N	-MET-	ASP-	ASP-	SER-	ALA-	PRO-	TYR-	ILE-
5'	-ATG	GAT-	GAT-	TCT-	GCA-	CCG-	TAT-	ATT-

9	10	11	12	13	14	15	16
GLY-	ALA-	ASN-	ASP-	ALA-	TRP-	ASP-	LEU-
GGA-	GCA-	AAT-	GAT-	GCA-	TGG-	GAT-	CTT-

17	18	19	20	21	22	23	24
GLY-	TYR-	THR-	GLY-	LYS-	GLY-	ILE-	LYS-
GGA-	TAT-	ACA-	GGA-	AAA-	GGA-	ATT-	AAA-

25	26	27	28	29	30	31	32
VAL-	ALA-	ILE-	ILE-	ASP-	THR-	GLY-	VAL-
GTT-							

33	34	35
GLU-	TYR-	ASN-

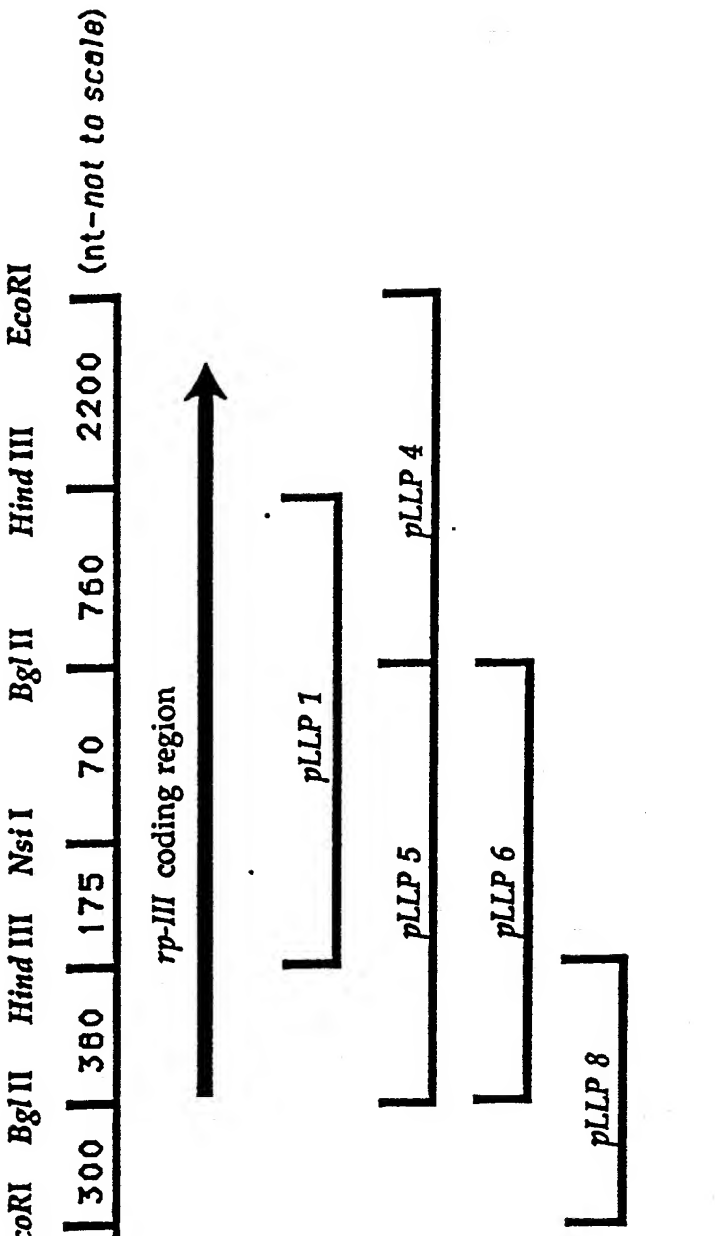
Nsi I
Hind III
EcoRI
Bgl II
Hind III
Nsi I
Bgl II
Hind III
EcoRI
1800
2000
300
380
175
70
760
2200
(nt-not to scale)
rp-III coding region
pLLP 1
pLLP 5
pLLP 4
pLLP 6
pLLP 8

Figure 3

```
001  ATC TTT CAC ATT TTT TCT AAA TAC AAA GGG GGA AAC ACA TTG AAA AAG GGG ATC ATT CGC
                                                      met lys lys gly ile ile arg

061  TTT CTG CTT GTA AGT TTC GTC TTA TTT TTT GCG TTA TCC ACA GGC ATT ACG GGC GTT CAG
     phe leu leu val ser phe val leu phe phe ala leu ser thr gly ile thr gly val gln

121  GCA GCT CCG GCT TCT TCA AAA ACG TCG GCT GAT CTG GAA AAA GCC GAG GTA TTC GGT GAT
     ala ala pro ala ser ser lys thr ser ala asp leu glu lys ala glu val phe gly asp

181  ATC GAT ATG ACG ACA AGC AAA AAA ACA ACC GTT ATA GTG GAA TTA AAA GAA AAA TCC TTG
     ile asp met thr thr ser lys lys thr thr val ile val glu leu lys glu lys ser leu

241  GCA GAA GCG AAG GAA GCG GGA GAA AGC CAA TCG AAA AGC AAG CTG AAA ACC GCT CGC ACC
     ala glu ala lys glu ala gly glu ser gln ser lys ser lys leu lys thr ala arg thr

301  AAA GCA AAA AAC AAA GCA ATC AAA GCA GTG AAA AAC GGA AAA GTA AAC CGG GAA TAT GAG
     lys ala lys asn lys ala ile lys ala val lys asn gly lys val asn arg glu tyr glu

361  CAG GTA TTC TCA GGC TTC TCT ATG AAG CTT CCA GCT AAT GAG ATT CCA AAA CTT CTA GCG
     gln val phe ser gly phe ser met lys leu pro ala asn glu ile pro lys leu leu ala

421  GTA AAA GAC GTT AAG GCA GTG TAC CCG AAC GTC ACA TAT AAA ACA GAC AAT ATG AAG GAT
     val lys asp val lys ala val tyr pro asn val thr tyr lys thr asp asn met lys asp

481  AAA GAC GTC ACA ATC TCC GAA GAC GCC GTA TCT CCG CAA ATG GAT GAC AGT GCG CCT TAT
     lys asp val thr ile ser glu asp ala val ser pro gln met asp asp ser ala pro tyr

541  ATC GGA GCA AAC GAT GCA TGG GAT TTA GGC TAC ACA GGA AAA GGC ATC AAG GTG GCG ATT
     ile gly ala asn asp ala trp asp leu gly tyr thr gly lys gly ile lys val ala ile

601  ATT GAC ACT GGG GTT GAA TAC AAT CAC CCA GAT CTG AAG AAA AAC TTT GGA CAA TAT AAA
     ile asp thr gly val glu tyr asn his pro asp leu lys lys asn phe gly gln tyr lys

661  GGA TAC GAT TTT GTG GAC AAT GAT TAC GAT CCA AAA GAA ACA CCA ACC GGC GAT CCG AGG
     gly tyr asp phe val asp asn asp tyr asp pro lys glu thr pro thr gly asp pro arg

721  GGC GAG GCA ACT GAC CAT GGC ACA CAC GTA GCC GGA ACT GTG GCT GCA AAC GGA ACG ATT
     gly glu ala thr asp his gly thr his val ala gly thr val ala ala asn gly thr ile

781  AAA GGC GTA GCG CCT GAT GCC ACA CTT CTT GCT TAT CGT GTG TTA GGG CCT GGC GGA AGC
     lys gly val ala pro asp ala thr leu leu ala tyr arg val leu gly pro gly gly ser

841  GGC ACA ACG GAA AAC GTC ATC GCG GGC GTG GAA CGT GCA GTG CAG GAC GGG GCA GAT GTG
     gly thr thr glu asn val ile ala gly val glu arg ala val gln asp gly ala asp val

901  ATG AAC CTG TCT CTC GGA AAC TCT TTA AAC AAC CCG GAC TGG GCG ACA AGC ACA GCG CTT
     met asn leu ser leu gly asn ser leu asn asn pro asp trp ala thr ser thr ala leu

961  GAC TGG GCC ATG TCA GAA GGC GTT GTC GCT GTT ACC TCA AAC GGC AAC AGC GGA CCG AAC
     asp trp ala met ser glu gly val val ala val thr ser asn gly asn ser gly pro asn

1021 GGC TGG ACA GTC GGA TCG CCG GGC ACA TCA AGA GAA GCG ATT TCT GTC GGT GCG ACT CAG
     gly trp thr val gly ser pro gly thr ser arg glu ala ile ser val gly ala thr gln

1081 CTG CCG CTC AAT GAA TAC GCC GTC ACT TTC GGC TCC TAC TCT TCA GCA AAA GTG ATG GGC
     leu pro leu asn glu tyr ala val thr phe gly ser tyr ser ser ala lys val met gly
```

```
1141  TAC AAC AAA GAG GAC GAC GTC AAA GCG CTC AAT AAC AAA GAA GTT GAG CTT GTC GAA GCG
      tyr asn lys glu asp asp val lys ala leu asn asn lys glu val glu leu val glu ala

1201  GGA ATC GGC GAA GCA AAG GAT TTT GAA GGG AAA GAC CTG ACA GGC AAA GTC GCC GTT GTC
      gly ile gly glu ala lys asp phe glu gly lys asp leu thr gly lys val ala val val

1261  AAA CGA GGC AGC ATT GCA TTT GTG GAT AAA GCG GAT AAC GCT AAA AAA GCC GGT GCA ATC
      lys arg gly ser ile ala phe val asp lys ala asp asn ala lys lys ala gly ala ile

1321  GGC ATG GTT GTG TAT AAC AAC CTC TCT GGA GAA ATT GAA GCC AAT GTG CCA GGC ATG TCT
      gly met val val tyr asn asn leu ser gly glu ile glu ala asn val pro gly met ser

1381  GTC CCA ACG ATT AAG CTT TCA TTA GAA GAC GGC GAA AAA CTC GTC AGC GCC CTG AAA GCT
      val pro thr ile lys leu ser leu glu asp gly glu lys leu val ser ala leu lys ala

1441  GGT GAG ACA AAA ACA ACA TTC AAG TTG ACG GTC TCA AAA GCG CTC GGT GAA CAA GTC GCT
      gly glu thr lys thr thr phe lys leu thr val ser lys ala leu gly glu gln val ala

1501  GAT TTC TCA TCA CGC GGC CCT GTT ATG GAT ACG TGG ATG ATT AAG CCT GAT ATT TCC GCG
      asp phe ser ser arg gly pro val met asp thr trp met ile lys pro asp ile ser ala

1561  CCA GGG GTC AAT ATC GTG AGC ACG ATC CCA ACA CAC GAT CCT GAC CAT CCA TAC GGC TAC
      pro gly val asn ile val ser thr ile pro thr his asp pro asp his pro tyr gly tyr

1621  GGA TCA AAA CAA GGA ACA AGC ATG GCA TCG CCT CAT ATT GCC GGA GCG GTT GCC GTT ATT
      gly ser lys gln gly thr ser met ala ser pro his ile ala gly ala val ala val ile

1681  AAA CAA GCC AAA CCA AAG TGG AGC GTT GAA CAG ATT AAA GCC GCC ATC ATG AAT ACC GCT
      lys gln ala lys pro lys trp ser val glu gln ile lys ala ala ile met asn thr ala

1741  GTC ACT TTA AAG GAT AGC GAT GGG GAA GTA TAT CCG CAT AAC GCT CAA GGC GCA GGC AGC
      val thr leu lys asp ser asp gly glu val tyr pro his asn ala gln gly ala gly ser

1801  GCA AGA ATT ATG AAC GCA ATC AAA GCC GAT TCG CTC GTC TCA CCT GGA AGC TAT TCA TAC
      ala arg ile met asn ala ile lys ala asp ser leu val ser pro gly ser tyr ser tyr

1861  GGC ACG TTC TTG AAG GAA AAC GGA AAC GAA ACA AAA AAT GAA ACG TTT ACG ATT GAA AAT
      gly thr phe leu lys glu asn gly asn glu thr lys asn glu thr phe thr ile glu asn

1921  CAA TCT TCC ATT AGA AAG TCA TAC ACA CTT GAA TAC TCA TTT AAT GGC AGC GGC ATT TCC
      gln ser ser ile arg lys ser tyr thr leu glu tyr ser phe asn gly ser gly ile ser

1981  ACA TCC GGC ACA AGC CGT GTT GTG ATT CCG GCA CAT CAA ACC GGG AAA GCC ACT GCA AAA
      thr ser gly thr ser arg val val ile pro ala his gln thr gly lys ala thr ala lys

2041  GTA AAG GTC AAT ACG AAG AAA ACA AAA GCT GGC ACC TAT GAA GGA ACG GTT ATC GTC AGA
      val lys val asn thr lys lys thr lys ala gly thr tyr glu gly thr val ile val arg

2101  GAA GGC GGA AAA ACG GTC GCT AAG GTA CCT ACA TTG CTG ATT GTG AAA GAG CCC GAT TAT
      glu gly gly lys thr val ala lys val pro thr leu leu ile val lys glu pro asp tyr

2161  CCG AGA GTC ACA TCT GTC TCT GTC AGC GAA GGG TCT GTA CAA GGT ACC TAT CAA ATT GAA
      pro arg val thr ser val ser val ser glu gly ser val gln gly thr tyr gln ile glu

2221  ACC TAC CTT CCT GCG GGA GCG GAA GAG CTG GCG TTC CTC GTC TAT GAC AGC AAC CTT GAT
      thr tyr leu pro ala gly ala glu glu leu ala phe leu val tyr asp ser asn leu asp
```

```
2281  TTC GCA GGC CAA GCC GGC ATT TAT AAA AAC CAA GAT AAA GGT TAC CAG TAC TTT GAC TGG
      phe ala gly gln ala gly ile tyr lys asn gln asp lys gly tyr gln tyr phe asp trp

2341  GAC GGC ACG ATT AAT GGC GGA ACC AAA CTT CCG GCC GGA GAG TAT TAC TTG CTC GCA TAT
      asp gly thr ile asn gly gly thr lys leu pro ala gly glu tyr tyr leu leu ala tyr

2401  GCC GCG AAC AAA GGC AAG TCA AGC CAG GTT TTG ACC GAA GAA CCT TTC ACT GTT GAA TAA
      ala ala asn lys gly lys ser ser gln val leu thr glu glu pro phe thr val glu OCH

2461  GAAAAAGCCCTGCCGATTCGGCAGGGCTTTTTAAAGATCAGTCAGCAAACGCCTCCTGCAATAAGCGATACG
```

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/01598

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C12P 21/02
US CL : 435/69.1, 219

II. FIELDS SEARCHED

Minimum Documentation Searched[4]

Classification System	Classification Symbols
U.S.	435/69.1, 219

Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched[5]

BIOSIS, MEDLINE, WPI, APS, JPABS, EMBL, GENBANK, UMBEL
SEARCH TERMS; PROTEASE, REDUCED, BACILLUS, FIGURE 4

III. DOCUMENTS CONSIDERED TO BE RELEVANT[14]

Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages[17]	Relevant to Claim No.[18]
Y	US, A, 4,828,994 (FAHNESTOCK ET AL) 09 MAY 1989, SEE ENTIRE DOCUMENT	1-21
Y	Journal of Bacteriology, Volume 160, No.1, issued October 1984, M. Y. Yang et al, "Cloning of the Neutral Protease Gene of Bacillus subtilis and the Use of the Cloned Gene to Create an In Vitro-Derived Deletion Mutation", pages 15-21, see entire document.	1-26
Y	EP, A, 0,257,189 (UDAKA ET AL) 02 MARCH 1988, SEE ENTIRE DOCUMENT.	1-21
Y	EP, A, 0,369,817 (SLOMA ET AL) 23 MARCH 1990, SEE ENTIRE DOCUMENT.	1-21
Y	WO, A, 86/01825 (FAHNESTOCK ET AL) 27 MARCH 1986, SEE ENTIRE DOCUMENT.	1-21
Y	US, A, 4,946,789 (UDEKA ET AL) 07 AUGUST 1990, SEE ENTIRE DOCUMENT.	1-21
Y	JOURNAL OF BACTERIOLOGY, VOLUME 158, NO.2, ISSUED MAY 1984, M. L. STAHL ET AL, "REPLACEMENT OF THE BACILLUS SUBTILISIN STRUCTURAL GENE WITH AN IN VITRO-DERIVED DELETION MUTATION", PAGES 411-418, SEE ENTIRE DOCUMENT.	1-26

* Special categories of cited documents:[15]

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search[2]	Date of Mailing of this International Search Report[2]
14 JUNE 1992	29 JUN 1992
International Searching Authority[1]	Signature of Authorized Officer[20]
ISA/US	DAVID B. SCHMICKEL

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

Y	EP, A, 0,227,260(SLOMA ET AL) 01 JULY 1987, SEE ENTIRE DOCUMENT.	1-26
Y	JOURNAL OF BACTERIOLOGY, VOLUME 172, NO. 2, ISSUED FEBRUARY 1990, A. SLOMA ET AL.,"GENE ENCODING A NOVEL EXTRACELLULAR METALLOPROTEASE IN BACILLUS SUBTILIS", PAGES 1024-1029, SEE ENTIRE DOCUMENT.	1-26

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE[1]

1. ☐ Claim numbers _, because they relate to subject matter (1) not required to be searched by this Authority, namely:

2. ☐ Claim numbers _, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out (1), specifically:

3. ☐ Claim numbers _, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING[2]

This International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Search Authority did not invite payment of any additional fee.

Remark on protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.